SAF/AAD Directory Structure and File Naming Conventions

Directorate of Departmental Publishing (SAF)

Capt Harris/SAF/AADXS/27 Jun 97/mch

07/12/97

Purpose

- Inform key AF publishing personnel of pending changes in SAF/AAD's directory structure (DS) and file naming convention (FNC)

2

07/11/97

Background Overview

The Beast

Today's Situation

- 5 distinct AAD file systems
- Unique directory structures
- Unnecessary variability is confusing
- Version control problems
- Automation scripting difficult
- Impossible to capture unique description of file in its name alone

07/11/97

SAF/AAD File Systems

System	Description
AFEPL Source	*Air Force Electronic Publications Library* CD-ROM
AFPDL Source (Solaris Unix)	*Air Force Publications Distribution Library* source files
AFPDL (Solaris Unix)	*Air Force Publications Distribution Library* FTP site
AFPDL BBS (DOS)	*AFPDL Bulletin Board System* for modem access (retired June 1997)
AFPUBS (BSD Unix)	*Air Force Publications* web site
Complete	Current local area network library of newly completed publications
Final	Current local area network library of newly completed data collection tools (forms)

Problems

- Manual directory and file management
  - Version control incidents
  - Not user friendly
  - Constant maintenance and troubleshooting
- Automation frustrations
  - Must create and maintain complex scripts to convert among structures
  - Any change must be manually rippled to each system

Problems (continued)

- No file naming scheme offers consistent, descriptive, unique file names within 8.3 convention

How Did We Get Here?

- Current systems evolved organically and independently (no unified planning)
- Users built download processes (some automated) around our structures
- CD-ROM and WWW were not factors
- New EP processes highlight problems

07/11/97

Goal

- Devise a single, consistent DS and FNC for AAD publications and forms repositories and distribution systems

07/11/97

Objectives of Changes

- Manage via an enterprise database management system (DBMS)
- Optimize for file management automation
- Reduce nonpurposeful redundancies
- Leverage unique pubs and forms numbers
- Optimize for online use
- Accommodate stand-alone use

07/11/97

Part 1 - Assumptions

07/11/97

1: Version Control is Critical

■ It is necessary to reduce opportunities for version control incidents to a minimum

■ Database Management and other automation will be key

07/11/97

2: Structured Approach is Necessary

- Directory structures that evolved organically
  - diverged from each other over time
  - fossilized and no longer meet requirements
  - became inflexible
- A systematic restructuring will be
  - consistent across repositories
  - optimized for new processes

07/11/97

3: Database Management is Key to Success

- Effective file management will treat files as elements in a relational or object database

- DBMS will track file versions, check-out, and check-in

07/11/97

4: DS and FNC will Facilitate Easier Automation Scripting

- Disparities among repositories necessitate many tedious conversions
- A unified convention will simplify automation efforts

07/11/97

5: The "AFEPL" Directory is Unnecessary

- A redundant folder
- Adds no additional useful information

07/11/97

6: DS Must Facilitate Online and CD-ROM Interaction

- Modes of use now leverage multiple digital distribution avenues
- DS should be the same across modes

07/11/97

7: Compliance with ISO 9660 Levels Unnecessary

- ISO 9660 specifies Max. 8-level DS
- 9 or more levels have no negative effect
- Should use as few levels as possible, for efficiency, but more if justified

07/11/97

8: Not All MAJCOMs / FOAs Have 32-bit OS

- Therefore we cannot have a FNC that depends on long file names in distribution repositories

9: You Can't Fully Describe with an 8.3 Filename

- Example of Problem:
  - \<pubtype\>\<series\>_\<pub#\>v\<vol#\>p\<part#\>c\<chap#\>s\<sec#\>i\<ic#\>.\<ext\>
  - AFMAN36_2020v11p10c10s12.pdf (24 characters + 3)
- You can't avoid moving some of this info to the folder names

07/11/97

10: Files will be Identified by both Folder and File Names

- Current ad-hoc file naming uses directory folder names to capture some essential file description information
- Any viable new scheme will simply formalize combining folder and file names for complete identifying info
- Will keep files in their folders when distributing

07/11/97

11: Interim Changes may Vary from Rule

- ICs are short-lived
- More likely to be needed case-by-case, rather than as part of a mass download
- Must be integrated into basic pub rapidly
- Probably need "IC" in filename

07/11/97

12: Each MAJCOM / FOA Evolved Unique DS & FNC

- In the absence of early central planning, each site evolved uniquely
- Each publisher's processes tie into their unique DS and FNC
- It's unlikely any standard approach will preclude the need for modifications at each site
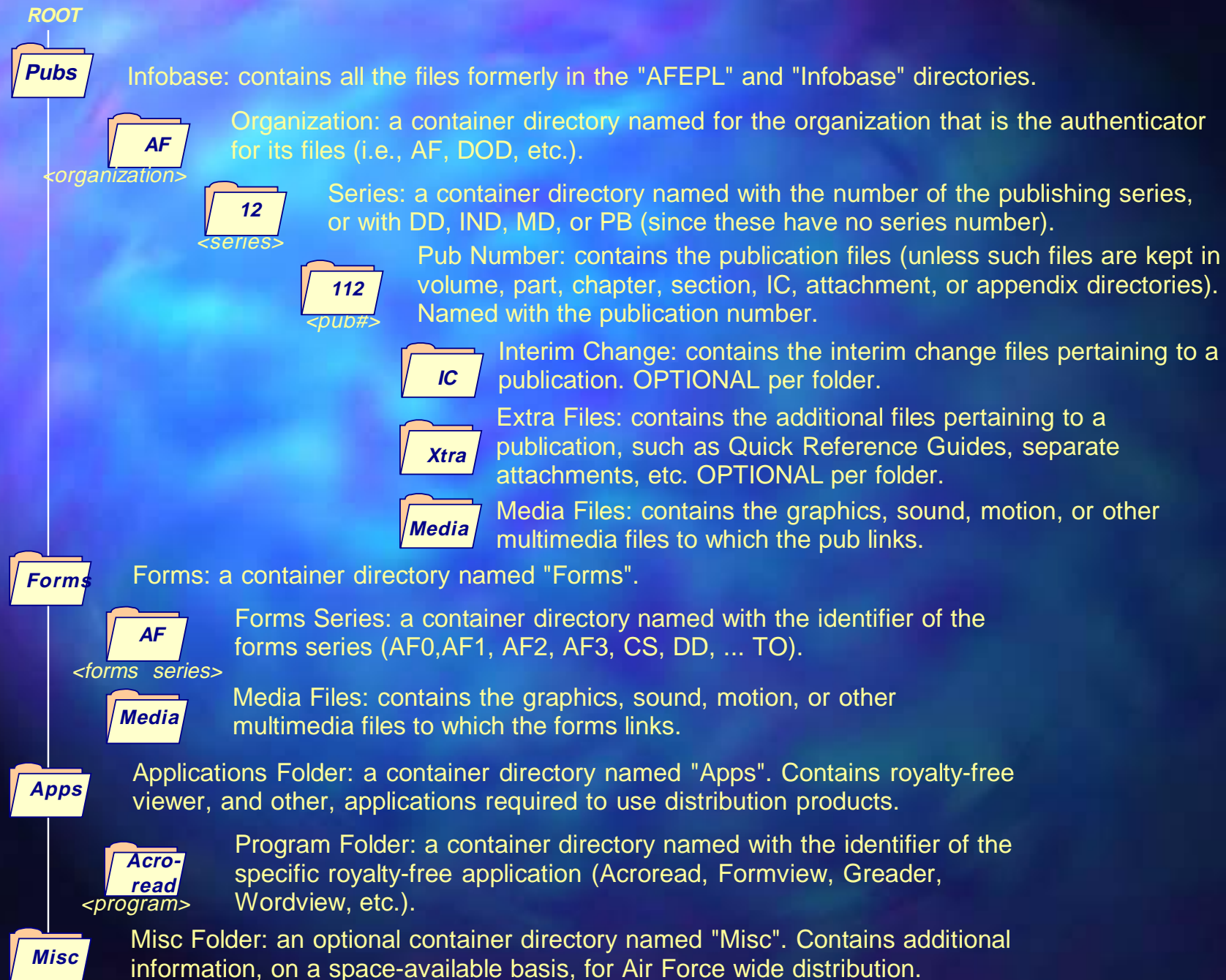- Everyone will have to adapt

07/11/97

Planned Approach

- AAD will develop a DS and FNC concordant with the assumptions
- AAD will coordinate implementation with MAJCOMs and FOAs
- AAD will allow leadtime (2 weeks) for customers to rewrite download scripts
- Incremental mirroring will alleviate version control or scripting difficulties

07/11/97

Part 2 - Proposed Directory Structure

07/11/97

SAF/AAD Distribution Systems Common Directory Structure

ROOT

Pubs — Infobase: contains all the files formerly in the "AFEPL" and "Infobase" directories.

AF
<organization> — Organization: a container directory named for the organization that is the authenticator for its files (i.e., AF, DOD, etc.).

12
<series> — Series: a container directory named with the number of the publishing series, or with DD, IND, MD, or PB (since these have no series number).

112
<pub#> — Pub Number: contains the publication files (unless such files are kept in volume, part, chapter, section, IC, attachment, or appendix directories). Named with the publication number.

IC — Interim Change: contains the interim change files pertaining to a publication. OPTIONAL per folder.

Xtra — Extra Files: contains the additional files pertaining to a publication, such as Quick Reference Guides, separate attachments, etc. OPTIONAL per folder.

Media — Media Files: contains the graphics, sound, motion, or other multimedia files to which the pub links.

Forms — Forms: a container directory named "Forms".

AF
<forms series> — Forms Series: a container directory named with the identifier of the forms series (AF0,AF1, AF2, AF3, CS, DD, ... TO).

Media — Media Files: contains the graphics, sound, motion, or other multimedia files to which the forms links.

Apps — Applications Folder: a container directory named "Apps". Contains royalty-free viewer, and other, applications required to use distribution products.

Acro-read
<program> — Program Folder: a container directory named with the identifier of the specific royalty-free application (Acroread, Formview, Greader, Wordview, etc.).

Misc — Misc Folder: an optional container directory named "Misc". Contains additional information, on a space-available basis, for Air Force wide distribution.

Publications Folder Names

- "Pubs" (literal; folder contains directory structure formerly kept in the "AFEPL" and "Infobase" folders)

- Organization Identifier (variable; e.g., "AF", "DoD", etc.)

- Series Identifier (variable two-character alphabetic or numeric series indicator; e.g., "PB", "MD", "01", "36", etc.)

07/11/97

Pubs Folder Names (cont.)

- Publication Number (four-digit number, e.g., "2903", "0001", "0210", etc.)

  NOTE: Use leading zeros, if necessary, to fill up the four-digits. This will cause folders to automatically sort in their proper sequence.

- "IC" (literal; stands for "interim change").

07/11/97

Pubs Folder Names (cont.)

- "Xtra" (literal; contains separately published files associated with a pub, such as quick reference guides; *does not apply to volumes*)
- "Media" (literal; contains graphics, sounds, motion, or other multimedia files to which the pub links)

Forms Folder Names

- "Forms" (literal: contains forms series and "Media" subfolders)

- Forms Series Identifier (variable; subfolder; e.g., AF0, AF1, AF2, AF3, CS, DD, DSP, GPO, Misc, OF, OMB, SF, TO, etc.)

- "Media" (literal; subfolder; contains graphics, sounds, and other multimedia files to which forms link)

07/11/97

Apps Folder Names

- "Apps" (literal; contains specific program subfolders)
- Program Folders (variable; subfolder; e.g., Acroread, Formview, Greader, Wordview, etc.)
  - May contain subfolders peculiar to the particular programs

07/11/97

Part 3 - File Naming

Understandable and Comprehensive

Pubs File Naming Conventions

- SAF/AAD will implement a structured file naming convention. Publication source and output files will be consistently named according to a single convention. There are three alternative conventions:

1. Publication date
2. 6+~1.3 (long name truncation)
3. 8.3

07/11/97

Option 1: Publication Date FNC

- Each publication could be named according to the syntax **<ddmmyyyy.aaa>**, where **dd** is the two-digit day-of-month, **mm** is the two-digit month, **yyyy** is the four-digit century, decade and year, and **aaa** is the three-digit file type extension.

07/11/97

Pub Date FNC Examples

- 01071997.pdf   1 Jul 1997 (PDF)
- 01092007.sgm   1 Sep 2007 (SGML)

NOTE 1: Redundant file names are permitted, but not in the same folder. It is the hierarchical path of folder names that adds context to the file name.

Note 2: A date as a file name is fine for computers and scripts, but is not intuitive for people downloading individual files (as opposed to sites that mirror our server). See options 2 and 3 for more intuitive names).

Option 2: 6+~1.3 Truncation FNC

- Each publication file would be named in such a way as to provide a fully descriptive long file name, on 32-bit operating systems, and a partially descriptive short 6+~1.3 file name on 16-bit operating systems

Truncation Examples

Long File Name	8.3 Truncated File Name
491000 AFPAM 49_1000 V6.pdf	491000~1.pdf
491000 AFPAM 49_1000 V7.pdf	491000~2.pdf
97IC02 AFMAN 49_2001 V12 P13 C34 S10 IC 97-2.pdf	97IC02~1.pdf
362001 AFPAM 36_2001	362001~1.pdf

When truncating a long file name, the 16-bit OS preserves the first 6 characters of the long name, adds a tilde (~), and assigns a number indicating the download sequence for files sharing the same first 6-characters. This is intriguing, but problematic.

These examples demonstrate the impossibility of achieving fully descriptive, unique file names within the 6+~1.3 name space.

07/11/97

8.3 FNC Alternative

- Each publication file would be named according to the syntax, <SSPPPPVV.ext>, where "SS" is the series number; "PPPP" is the publication number; "VV" is the number of any part of a publication; such as a volume, that is stored as a separate file; and "ext" is the default file extension.

8.3 FNC Alternative Examples

- 36290301.pdf   (AFI 36-2903)
- 33036009.sgm   (AFI 33-360 V.9)
- 36000101.pdf   (AFPD 36-1)
- 10022211.sgm   (AFH 10-222 V.11)

NOTE: Use leading zeros to make each segment of the number occupy its full space.

07/11/97

Comparison of Alternatives

- Neither consistently creates unique, fully descriptive 8.3 file names.

- Both 8.3 alternatives depend on folder paths for complete description of files.

- Only the second (truncation) provides unique, fully descriptive *long* file names.

- On download, 16-bit systems may need to rename the truncated file.

- 8.3 FNC is the most descriptive within our current constraints. It is AAD's choice.

07/11/97

Misc. File Names

- Files below the Publication Number level may include those stored in the "IC", "Media" and "Xtra" folders. These will follow a descriptive name convention incorporating the elements listed on the following three slides. These elements will not be used in publication names.

Misc. File Name Elements - 1

- Element Type Identifier
  - "F" for figure (graphic in pub)
  - "G" for graphic in form
  - "T" for table (a graphically captured image of a table)
  - "S" for sound (audio)
  - "M" for motion (animation or video) elements

07/11/97

Misc. File Name Elements - 2

- Attachment Identifier ("A" — to associate a media file with an attachment)
- Attachment Number (a two-digit integer; leading zero)
- Paragraph Number Identifier (literal: "P")

Misc. File Name Elements - 3

- Chapter Number (a two-digit integer; leading zero)

- Underscore (literal)

- Paragraph Number (a three-digit integer; leading zero)

Long File Name Elements - 4

- Period (delimiter between file name and file type extension)

- File Type Extension (one to three characters; application defaults)

- Space (literal; long file names only)

Misc. File Name Examples

File Name	Description
F4_3.jpg	Fig. 4-3
F1T3_5.gif	Fig. 1 in Table 3-5
F3A2.jpg	Fig. 3 in Atch. 2
F2_3A4.gif	Fig. 2-3 in Atch. 4
T1_1.gif	Table 1-1 (captured image)
M2A4.avi	Motion clip 2 in Atch. 4

Part 4: SAF/AAD Preferred FNC Option

8.3 Pubs Naming

- Simple
- Facilitates automation
- Easy to interpret (by humans or machines)
- With a few subfolders, handles all pubs types
- Can be implemented rapidly

Forms File Naming

NO CHANGE: Forms file naming does not follow a meaningful scheme, as pubs do. They will continue to follow current FNC, with the exception that FormFlow 2 files will have different extensions than PerForm Flow files.

Part 5: Implementation

Step 1: Notify Customers

- MAJCOM and FOA Publishers

- PDOs

- End users

Step 2: AAD Implementation

- Build new DS on separate systems (1 to 2 weeks)
- Copy and rename files to new systems (2 weeks)
- MAJCOMS / FOAs / PDOs mirror AFPDL
- Provide brief dual-system adjustment period (two-weeks)
- Retire old structures

Additional Notes

AFPDL

- Evolved far beyond its main role of source and output file transfers
- Will be reincarnated in NT 4
- Will have the same library structure as AFPUBS and AFEPL
- Will be kept free of extraneous folders and files
- *A separate FTP site will provide misc. publishing information

Comments

Send comments to AFPIC or to
mcharris@Air Force.pentagon.mil

Capt Mark Harris, SAF/AADXS
DSN: 754-2441